

THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

UNCLASSIFIED

A 895116

Armed Services Technical Information Agency

Reproduced by
DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON, 2, OHIO

This document is the property of the United States Government. It is furnished for the duration of the contract and shall be returned when no longer required, or upon recall by ASTIA to the following address: Armed Services Technical Information Agency, Document Service Center, Knott Building, Dayton 2, Ohio.

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

UNCLASSIFIED

O. N. R. Research Memorandum No. 31
(Preliminary)

FC
BAC

AGGREGATION OF VARIABLES IN DYNAMICAL SYSTEMS

By

Herbert A. Simon*

July 1955

Graduate School of Industrial Administration

Carnegie Institute of Technology

*Research undertaken for the project Planning and Control of Industrial Operations, under contract with the Office of Naval Research. The paper is released for limited circulation in order to facilitate discussion, and invite criticism. It is tentative in character and should not be referred to in publications without the permission of the author(s).

Reproduction of this manuscript in whole or in part is permitted for any purpose of the United States Government. Contract N ONR - 76001, Project NR 047001.

AD No. 29 3516
ASTIA FILE COPY

**Best
Available
Copy**

AGGREGATION OF VARIABLES IN DYNAMICAL SYSTEMS

Apologia

Dittoed manuscripts are by definition preliminary. This one is especially so, and I am under no illusions as to the rigorousness of the proofs in Sections 4 and 5. On the other hand, the topic dealt with here has to do fundamentally with approximations, and I have a notion that in this instance heuristic is of more importance than rigor. Moreover, if anyone wishes to make matters more rigorous the proofs provided here do, I believe, mark a clear route; and patience and a liberal sprinkling of epsilons will fill in the details.

Acknowledgment: I have had numerous very enlightening discussions with Allen Newell on the subject of this paper.

1. Introduction

The need to deal with many problems of economic theory in terms of aggregates is obvious. The complete Walrasian system, and the more modern dynamic embellishments of it are relatively barren of results for macroeconomics and economic policy. Hence, by sheer necessity, we use highly aggregated systems, often without having much more than the same necessity as a justification for their use. Perhaps the most important result to date in the attempt to justify aggregation under certain circumstances is the Lange-Hicks condition, about which we shall say more later.

The development of the Leontief input-output model, and the concern with actual numerical coefficients in rather highly disaggregated versions of this model have renewed interest in aggregation.

There was at first, perhaps, the fond hope that modern computing equipment would handle matrices of any size we might be likely to concern ourselves with, and hence would obviate the need for aggregation. By now, it is clear that the size of the models we should like to handle has outstripped the development of computers—and is likely to continue to do so as long as the time required to invert a matrix increases with the cube of the number of rows and columns. A number of experiments have already been made (e.g., Whittin and Morgenstern) to determine whether aggregation can be used to obtain approximate inverses more economically.

Most discussions of aggregation have taken place in the context of an algebraic model. The problem may be stated in general terms as follows: We have two sets (not necessarily distinct) of variables— $\{x_i\}$ and $\{y_i\}$ —and a system of equations giving the y 's as functions of the x 's. We wish to know under what circumstances there exists a function, $X(x_1, \dots, x_n)$, of the x 's and another function, $Y(y_1, \dots, y_n)$, of the y 's, such that a relation, $Y = \underline{g}(X)$, between Y and X can be deduced from the given system of equations relating the x 's and y 's. Sometimes additional conditions are imposed—e.g., the functions Y and X are given, or it is required that the relation \underline{g} possess certain properties. (For a comprehensive treatment, see Theil).

The conditions that must be satisfied in order for aggregation to be strictly possible are very severe. Whether these conditions are exactly satisfied in any practical situation is, of course, not an important question. We would be perfectly satisfied with aggregative models that gave us only approximate results—we have no illusions,

after all, that any model we might employ is more than an approximate description of reality. Our objective in exploring the aggregation problem should be to seek rules and criteria—exact or heuristic—that will give us clues as to what variables to aggregate, and will indicate to us under what circumstances aggregation is likely to yield satisfactory approximations.

The Lange-Hicks condition is a criterion of this kind. It states that if two or more variables always move together, then they may be aggregated into a single variable, which will be an appropriately weighed average of the original variables. This is a useful criterion, since it tells us that we may aggregate classes of commodities that are perfect substitutes—or that are approximately so.

At another level, the Lange-Hicks condition is unsatisfactory, for it requires that we know as a datum which variables move together. In actual fact, it may be part of our problem to discover this. We may be confronted with a dynamical system involving a large number of variables, and may have to infer from the equations of the system which variables will move together—or will behave nearly enough in this manner to warrant aggregation. This is the essential problem we shall set ourselves here: to determine conditions that, if satisfied by the equations of a (linear) dynamical system, will permit approximate aggregation of variables. Note that we will be interested in sufficient, rather than necessary, conditions. Hence, we may also view our task as that of discovering one or more classes of dynamical systems that permit aggregation of their variables.

2. Properties of "Nearly-Decomposable" Matrices

For purposes of exposition, it will be convenient to limit ourselves, for the most part, to stochastic matrices—that is, non matrices all of whose elements are non-negative and with the sum of the elements in each row equal to 1: $p_{ij} \geq 0$ ($i, j = 1, \dots, n$), and $\sum_j p_{ij} = 1$ ($j = 1, \dots, n$). The results may readily be generalized (Robert Solow, "On the Structure of Linear Models," *Econometrica*, January, 1952, pp. 42-43), but the restriction will perhaps make the intuitive basis of the analysis clearer than would a more general treatment. With this interpretation, the index i runs over the n possible states of the system, $p_i(t)$ represents the probability that the system is in the i^{th} state at time t , and p_{ij} represents the conditional probability that the system, if in state i at time t , will be in state j at time $(t+1)$.

Now suppose that by an approximate permutation of rows and corresponding columns the stochastic matrix $P^t = \left\| p^t_{ij} \right\|$ can be arranged as follows:

$$(2.1) \quad P^t = \left\| \begin{array}{cccc} P^t_1 & & & \\ & A^t_1 & & \\ & & P^t_2 & \\ & & & I \\ & & & & P^t_3 \\ & & & & & P^t_4 \\ & & & & & & P^t_N \end{array} \right\|$$

where the P^t_i are square submatrices and the remaining elements, not displayed, are all zero. Then the matrix is completely decomposable. Suppose further that none of the submatrices, P^t_i , are themselves decomposable.

Let the system, at time t_0 , be in the state I_1 , where I_1 is one

of the states belonging to the submatrix P_I^i . Then it is obvious that at all subsequent times the system will continue to be in one of the states belonging to P_I^i —that the probability of transition to any state outside this subset is zero. Further, if the initial position of the system is given by a probability distribution over the states belonging to P_I^i , all other states having zero probability, then the limiting distribution of the system, after the passage of an indefinite period of time, will be independent of this initial distribution. If, on the other hand, the initial distribution assigns non-zero probabilities to states belonging to more than one of the indecomposable submatrices, P_I^i , then the limiting distribution will not be independent of the initial distribution.

Let the initial distribution be $p_i(0)$, and define:

$$(2.2) \quad P_I(0) = \sum_{i \in I} p_i(0); \quad P_{I/I} = P_I/P_I$$

then the final distribution will be:

$$(2.3) \quad \bar{P}_I = P_I(0) \cdot \bar{P}_{I/I}$$

where the $\bar{P}_{I/I}$ (but not the \bar{P}_I) are independent of the initial distribution. The $\bar{P}_{I/I}$ are the characteristic vectors of the submatrix P_I^i corresponding to the largest characteristic root, $\lambda_1 = 1$. Further, the distribution of (2.3) is an equilibrium distribution in that, if this distribution is once attained by the system, it will not change over time.

Now, suppose the stochastic matrix to be slightly altered so that there is now at least one path with non-zero probability leading, in not more than n steps, from any one state of the system to any other state. That is to say, we suppose certain of the zero elements in P^i

that lie outside the submatrices, P_i^i , to be replaced by very small non-zero elements, and certain elements inside the P_i^i to be correspondingly reduced so as to preserve the row-sum condition. We will refer to the modified matrix, P , obtained from P^i in this fashion, as a "nearly-decomposable" matrix. The term "nearly-decomposable," like the term "very small non-zero elements", is not, of course, precise. We will be content to leave the criterion of smallness undefined. The smaller these elements are, the closer to the exact solution will be the approximation to be discussed here, but we shall not attempt to evaluate explicitly the goodness of the approximations.

Comparing the original decomposable matrix, P^i , and the modified, nearly-decomposable matrix, P , we may readily reach the following conclusions, which depend on the fact that the characteristic roots of a matrix are continuous functions of the matrix elements:

1. Unity will be the largest characteristic root of both P^i and P . It will be a root of multiplicity N of P^i (since each indecomposable submatrix of P^i will have unity as a simple root); but a simple root of P .
2. The matrix P^i will have $(n-N)$ roots that are less than 1 in absolute value; the matrix P will have a set of $(n-N)$ roots in one-one correspondence to these—the difference between corresponding roots of P^i and P being a continuous function of the differences between the two coefficients of the matrices.
3. The matrix P will have $(N-1)$ roots very close to unity. We will suppose the difference between P^i and P to be sufficiently small

that the difference between unity and the smallest of these roots is very much smaller (e.g., of the order of one percent) than the difference between unity and the largest of the remaining roots of P .

This isomorphism between the characteristic roots of P^i and the roots of P allows us to make a qualitative analysis of the structure of the latter system. The roots of P belong to three classes:

(1) the simple root at unity; (2) the $(N-1)$ next largest roots (corresponding to the remaining roots of P^i at unity); and (3) the $(n-N)$ smaller roots. The latter roots can, in turn, be partitioned into subsets corresponding to the submatrices of P^i .

We see that the roots of class (3) are involved in the process whereby equilibrium is reached in the distribution within each of the indecomposable subsets of states of the system. The roots of class (2) are involved in the process of attaining equilibrium among subsets. The largest root is associated with the steady-state distribution of the entire system.

The smaller a characteristic root, the more quickly is the transient damped with which this root is associated. We have assumed the difference between P^i and P to be so small that the roots of class (3) will be smaller than the roots of class (2). Hence, equilibrium within (the equilibrium among subsets. In first approximation, we may view) subsets will be attained more quickly than (the equilibrating process as proceeding in two successive steps:

(1) we reach the equilibrium represented by (2.3), corresponding to the equilibrium of the completely decomposable system for the same

initial distribution;

(2) maintaining this equilibrium within each subset at all times, we attain, through a slow change in the p_I , the general equilibrium of the engine system.

Notice that in the second stage of this process, all the p_i within any single subset vary proportionately—hence satisfy the Hicckelange condition for aggregation. That is, the long-range dynamic behavior of the system depends only on the p_I . As for the short-range dynamic behavior of the system, we may regard it as composed of N independent indecomposable parts—corresponding to the submatrices P_I —and may solve each of these subsystems independently of the others.

Let n_I be the number of rows (and columns) in P_I , and hence the number of states in the corresponding subset ($\sum n_I = n$). Then, by the process of decomposition just described, we have replaced a dynamical process represented by a $n \times n$ matrix with a superposition of dynamical processes represented by N matrices of size $n_I \times n_I$, ..., $n_N \times n_N$, respectively, and one additional matrix of size $N \times N$. Later, we shall show that this decomposition permits us to replace the task of inverting the matrix P with the much less laborious task of inverting the matrices of the decomposed system.

3. A Physical Illustration

Before we proceed with a more careful statement of the mathematics that underlies our analysis, it may be useful to provide an example of a physical system that can be, approximately, decomposed in the manner just described. The illustration will be useful, also, in showing that our conclusions do not depend on the stochastic properties of the matrix we have discussed thus far. We shall see that the

principle of aggregation we are employing is essentially that which justifies the replacement of microvariables by macrovariables in classical thermodynamics.

Let us consider an imaginary building, whose outside walls provide perfect thermal insulation from the environment. The building is divided into a large number of rooms, the walls between them being good, but not perfect, insulators. Each room is divided into a number of offices by partitions. The partitions are only poor insulators. A thermometer hangs in each of the offices.

Let us suppose that at time t_0 the various offices within the building are in a state of thermal disequilibrium—there is a wide variation in temperature from office to office and from room to room. If we take new temperature readings at time t_1 , several hours after t_0 , what will we find? At t_1 , there will be very little variation in temperature among the offices that are in each single room, but there may still be large temperature variations between rooms. If we take readings again at time t_2 , several days after t_1 , we may find that there is an almost uniform temperature throughout the building, the temperature differences between rooms having virtually disappeared.

The well-known equations for the diffusion of heat allow us to represent this situation by a system of differential equations—or approximately by a system of difference equations.

Let $T_{j,i}(t)$ be the temperature of the j^{th} office which is in the i^{th} room, at time t . Let $T(t)$ be the vector consisting of these temperatures as components, $T(t) = \{T_{1,1}, T_{1,2}, \dots, T_{1,j}, \dots, T_{N,k}, \dots, T_{N,m}\}$

Then we will have:

$$(3.1) \quad T(t+1) = A^T T(t)$$

where A is a matrix whose coefficients, a_{ij} , represent the rates of heat transfer between office i and office j per degree difference in temperature.

We wish now to represent the fact that a temperature equilibrium within each room is reached rather rapidly, while a temperature equilibrium among rooms is reached only slowly. This will be the case if the a_{ij} are generally large when i and j are offices in the same room, and if the a_{ij} are close to zero when i and j are offices in different rooms—that is to say, if the matrix A is nearly decomposable.

When this is the case, and as long as we are not interested in the high frequency fluctuations in temperature among offices in the same room, we can learn all we want to know about the dynamics of this system by placing a single thermometer in each room—it is unnecessary to place a thermometer in each office.

4. Time Paths of Nearly Decomposable Systems

We will return now to the special case of stochastic systems and examine the dynamical process in more detail. The system is given by:

$$(4.1) \quad p_j(t+1) = \sum_{i=1}^n p_i(t) a_{ij} \quad (j=1, \dots, n)$$

If we solve this system of difference equations to obtain the p_j 's as explicit functions of time, we get:

$$(4.2) \quad p_j(t) = \sum_{i=1}^n p_i \lambda_i^t \quad (j=1, \dots, n)$$

But we know also that:

$$(4.3) \quad p_j(t) = \sum_{i=1}^n p_i(0) p_{ij}^{(t)} \quad (j=1, \dots, n)$$

Now it is possible (provided that all the characteristic λ_p are distinct) to represent the right-hand side of (4.3) as a sum that is equal term by term to the right-hand side of (4.2). (see Wedderburn, pp. 25-6) To do this we proceed as follows:

For any $n \times n$ matrix, A , with n distinct characteristic roots, we can find a set of n idempotent matrices, $\alpha^{(p)}$, ($p=1, \dots, n$), with the following properties:

$$(i) \quad \alpha^{(p)} \cdot \alpha^{(p)} = \alpha^{(p)} \quad (\text{idempotence})$$

$$(ii) \quad \alpha^{(p)} \cdot \alpha^{(q)} = 0, \text{ for } p \neq q$$

$$(iii) \quad \sum_{p=1}^n \alpha^{(p)} = I \text{ where } I \text{ is the identity matrix, } I = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

$$(iv) \quad a_{ij} = \sum_{p=1}^n \lambda_p \alpha_{ij}^{(p)}$$

Now, consider the matrix A^t . Taking the t^{th} power of (iv) and using (i) and (ii), we find, for $A^t = \begin{bmatrix} a_{ij}^{(t)} \end{bmatrix}$

$$(v) \quad a_{ij}^{(t)} = \sum_{p=1}^n \lambda_p^t \alpha_{ij}^{(p)}$$

Using this representation, we may express $P^{(t)} = P^t$ thus:

$$(4.4) \quad p_{ij}^{(t)} = \sum_{p=1}^n \lambda_p^t \pi_{ij}^{(p)}$$

where, the $\pi^{(p)}$ are matrices satisfying (i), (ii) and (iii) above, substituting (4.4) in (4.3) and comparing the resulting expression with (4.2), we get:

$$(4.5) \quad a_{je} = \sum_{i=1}^n p_i(0) \pi_{ij}^{(e)}$$

Suppose now that $P = \begin{vmatrix} p_{ij} \end{vmatrix}$ is nearly decomposable, and let P' be the corresponding decomposable matrix obtained from P by an appropriate limiting process. By examination of the limiting process we can again divide the roots of P into three sets: (1) $\lambda_1 = 1$; (2) $\lambda_2, \dots, \lambda_N$ (corresponding to the remaining roots of unity in P'); (3) $\lambda_{N+1}, \dots, \lambda_n$.

Consider one of the indecomposable submatrices, P'_I , of P' and designate its roots by $\lambda_2^I = 1, \lambda_{N+1}^I, \dots, \lambda_{N+k}^I$. In exactly the same way as in (4.4) we will have:

$$(4.6) \quad p_I^t = \pi^t(I) + \sum_{\substack{\rho=N+1 \\ \rho \neq N+k}}^{N+k} \lambda_\rho^t \pi^t(\rho)$$

Next consider the $n \times n$ matrices that are formed from the above matrices by bordering them with an appropriate number of rows and columns having all elements zero. We will designate these bordered matrices with the same symbols as are employed for the corresponding $n_I \times n_I$ matrices in (4.6). It will be clear from the context in each instance which set of matrices is intended.

Now, we may write:

$$(4.7) \quad P^t = P_A^t + \dots + P_I^t + \dots + P_N^t = \sum_{I=A}^N \pi^t(I) + \sum_{\rho=N+1}^n \lambda_\rho^t \pi^t(\rho)$$

and similarly for the t^{th} power of P' :

$$(4.8) \quad p^t = \sum_{I=A}^N \pi^t(I) + \sum_{\rho=N+1}^n \lambda_\rho^t \pi^t(\rho)$$

Now, for P^t , we will have the corresponding representation:

$$(4.9) \quad p^t = \pi^t(1) + \sum_{\rho=2}^N \lambda_\rho^t \pi^t(\rho) + \sum_{\rho=N+1}^n \lambda_\rho^t \pi^t(\rho)$$

We observe that the last summation in (4.9) will be as nearly equal as we please, term by term, to the corresponding summation in

(4.8), — provided only that we take P sufficiently close to P' .

We will have also the approximate equality:

$$(4.10) \quad \sum_{I=A}^N \pi^*(I) \approx \pi(1) \approx \sum_{q=2}^N \pi^{(q)}$$

But in (4.10) the correspondence cannot be term-by-term, for each column of $\pi(1)$ has as elements the \bar{p}_i , while each column of $\pi^{(I)}$ has as elements the \bar{p}_i/I . If i and i' both belong to the subset I , then for moderate or large t , we have^{*}:

$$(4.11) \quad p_{ij'}^{(t)} \approx p_{ij'}^{(t)} \approx \bar{p}_{i'}/I$$

If i belongs to I and j to J , $J \neq I$, then we will have, approximately

$$(4.12) \quad p_{ij}^{(t)} \approx \sum_{a=0}^{t-1} \sum_{i' \in I} \sum_{j' \in J} p_{ii'}^{(t-a-1)} p_{i'j'} p_{j'j}^{(a)},$$

and using (4.11),

$$(4.13) \quad p_{ij}^{(t)} \approx \sum_a \sum_{i' \in I} \sum_{j' \in J} \bar{p}_{i'}/I p_{i'j'} p_{j'j} \\ \approx p_{j/J} p_{IJ}^{(t)}$$

$$\text{where } p_{IJ} \approx \sum_{i' \in I} \sum_{j' \in J} \bar{p}_{i'}/I p_{i'j'}.$$

Equation (4.12) assumes the $p_{i'j'}$ to be very small when i' and j' belong to different subsets. In that case, a transition from state i to state j in t periods may be assumed to involve a single transition from one subset to another — paths involving several such transitions

^{*} By "moderate or large t " we mean t sufficiently large that $\lambda \approx 0$ for $\lambda = \lambda_0, \dots, \lambda_N$; $\lambda_q \neq 0$ for $q=1, \dots, N$.

having a vanishingly small probability. Hence, on the right-hand side of (4.12), we have retained only the terms involving $p_{ij}^{(1)}$ raised to the first power. Similarly, for reasonably large t , the terms in which a is very small or very close to $(t-1)$ will be negligible compared with the sum of terms in which a is of the order of $t/2$. But for all the latter terms, (4.11) will hold, thus justifying the step leading to (4.13),

From (4.13) we see that $p_{ij}^{(t)}$ depends only on j , I , and J , and is independent of i . But for moderate or large t we also have, from (4.10)

$$(4.14) \quad \sum_{I=1}^N \pi_{ij}^{(I)} - \pi_{ij}^{(1)} = \bar{p}_{j/J} \delta_{IJ} - \bar{p}_j \\ = \bar{p}_{j/J} (\delta_{IJ} - \bar{p}_J) = \sum_{I=2}^N \pi_{ij}^{(I)}$$

and from (4.13) and (4.9)

$$(4.15) \quad p_{ij}^{(t)} - \pi_{ij}^{(1)} = \bar{p}_{j/J} p_{ij}^{(t)} - \bar{p}_j = \sum_{I=2}^N \lambda_I^t \pi_{ij}^{(I)}$$

But for (4.14) and (4.15) to hold identically in t , we must have:

$$(4.16) \quad \pi_{ij}^{(p)} = a_{IJ}^{(p)} [\bar{p}_{j/J} \delta_{IJ} - \bar{p}_j] \quad (p=2, \dots, N)$$

We have now reached the important result that for each of the first N idempotent matrices, $\pi^{(p)}$, elements within a given bloc (e.g. the bloc of rows $i \in I$ and columns $j \in J$) will vary only with $\bar{p}_{j/J}$, and will otherwise be functions only of I and J . This permits us, for purposes of inversion, to replace these matrices with aggregative matrices having

a single row and column for each subset I . The method of doing this will be described in the next section.

5. Aggregation and Inversion

From (4.4), we see that if we can express a matrix as a sum of idempotent components, we can readily invert it. For, if we take $t = -1$, we obtain immediately:

$$(5.1) \quad P^{-1} = P_{1j}^{-1} = \sum_{p=1}^{(-1)} \lambda_p^{-1} \pi_{1j}^{(p)} \quad (i, j = 1, \dots, n)$$

That is, to find P^{-1} , we multiply each idempotent matrix by the reciprocal of the corresponding characteristic root, and sum over all such matrices. Ordinarily, this does not provide us with a useful computational procedure, for the matrices, $\pi^{(p)}$, are difficult to compute. In the present case, however, we shall see that (5.1) leads us to a very convenient method of inversion that does not require, moreover, the explicit computation of the $\pi^{(p)}$.

Using (4.7) and (4.9), we have immediately:

$$(5.2) \quad P^{-1} = \sum_{p=1}^{(-1)} \lambda_p^{-1} \pi^{(p)} + \sum_{I=A}^N (P_I^1)^{-1} - \sum_{I=A}^N \pi^{(I)}$$

The last two terms of (5.2) are found by inverting the P_I^1 and computing the characteristic vectors $[\pi^{(I)}]$ associated with the roots at unity. It remains to find the term in square brackets.

Suppose we have any matrix, $A = \left\| a_{ij} \right\|$, such that:

$$(5.3) \quad a_{ij} = a_{j/J} a_{IJ} \quad (1 \leq i, j \leq J)$$

Then it follows that:

$$(5.4) \quad a_{1j}^{(t)} = a_{j/J} a_{1J}^{(t)} \quad (\text{for all } t)$$

We prove this by induction:

$$(5.5) \quad a_{1j}^{(2)} = \sum_{J'} \sum_{J''} a_{J'J''} a_{j'/J'} a_{1J''} a_{j/J} a_{J'J''} \\ = a_{j/J} \sum_{J'} a_{1J'} a_{J'J} = a_{j/J} a_{1J}^{(2)}$$

Similarly, if $a_{1j}^{(t-1)} = a_{j/J} a_{1J}^{(t-1)}$ we have:

$$(5.6) \quad a_{1j}^{(t)} = \sum_{J'} \sum_{J''} a_{J'J''} a_{j'/J'} a_{1J''}^{(t-1)} a_{j/J} a_{J'J''} \\ = a_{j/J} a_{1J}^{(t)}$$

In particular, it follows that $a_{1j}^{(-1)} = a_{j/J} a_{1J}^{(-1)}$. Now let A be equal to the sum in brackets in (5.2):

$$(5.7) \quad A = \pi^{(1)} + \sum_{p=2}^N \lambda_p^{-1} \pi^{(p)}$$

Then, from (4.11), (4.13), and (4.66), we know that A satisfies

(5.3), hence that:

$$(5.8) \quad a_{1j}^{(-1)} = \bar{p}_{j/J} p_{1J}^{(-1)}$$

where, as in (4.13)

$$(5.9) \quad p_{1J} = \sum_{i \in I} \sum_{j \in J} \bar{p}_{i/I} p_{ij}$$

We now have a practical procedure for inverting the nearly-decomposable matrix, P_{\circ} .

- (1) We find the inverses of the P_I^i .
- (2) We compute the $\bar{p}_{i/I}$ and form the aggregative matrix p_{1J} with elements given by (5.9)
- (3) We invert p_{1J} , and find $a_{1j}^{(-1)}$ from (5.8).
- (4) We now obtain P^{-1} directly from (5.2)

6. A Numerical Example

The proposed procedure for matrix inversion can be illustrated by a simple example. Consider the nearly-decomposable matrix:

$$(6.1) \quad P \approx \begin{vmatrix} .9700 & .0295 & .0005 & 0 \\ .0200 & .9800 & 0 & 0 \\ 0 & 0 & .9600 & .0400 \\ .0002 & .0002 & .0396 & .9600 \end{vmatrix}$$

We take as P' :

$$(6.2) \quad P' = \begin{vmatrix} .9700 & .0300 & 0 & 0 \\ .0200 & .9800 & 0 & 0 \\ 0 & 0 & .9600 & .0400 \\ 0 & 0 & .0400 & .9600 \end{vmatrix} = \begin{vmatrix} P_1 & 0 \\ 0 & P_2 \end{vmatrix}$$

We readily find:

$$(6.3) \quad P_1^{-1} = \begin{vmatrix} 1.031579 & -.031579 \\ -.021053 & 1.021053 \end{vmatrix} \quad P_{j/1} = \begin{vmatrix} .4 & .6 \end{vmatrix}$$

$$P_2^{-1} = \begin{vmatrix} 1.043478 & -.043478 \\ -.043478 & 1.043478 \end{vmatrix} \quad P_{j/2} = \begin{vmatrix} .5 & .5 \end{vmatrix}$$

We now compute p_{IJ} and its inverse:

$$(6.4) \quad \begin{vmatrix} p_{IJ} \end{vmatrix} = \begin{vmatrix} .9998 & .0002 \\ .0002 & .9998 \end{vmatrix}$$

$$(6.5) \quad \begin{vmatrix} p_{IJ} \end{vmatrix}^{-1} = \begin{vmatrix} 1.0002 & -.0002 \\ -.0002 & 1.0002 \end{vmatrix}$$

and, from (5.8):

$$(6.6) \quad \left\| \begin{matrix} (-1) \\ a_{ij} \end{matrix} \right\| = \begin{vmatrix} .40008 & .60012 & -.0001 & -.0001 \\ .40008 & .60012 & -.0001 & -.0001 \\ -.00008 & -.00012 & .5001 & .5001 \\ -.00008 & -.00012 & .5001 & .5001 \end{vmatrix}$$

From (6.3) and (6.6), using (5.2), we obtain P^{-1} :

$$(6.7) \quad P^{-1} \approx \begin{vmatrix} 1.031659 & -.031459 & -.0001 & -.0001 \\ -.020973 & 1.021173 & -.0001 & -.0001 \\ -.00008 & -.00012 & 1.043578 & -.043578 \\ -.00008 & -.00012 & -.043378 & 1.043578 \end{vmatrix}$$

The same numerical example provides us with an insight into the dynamical process as expressed in (4.9). From (6.1), we compute P^{128} and $P^{(128)^2}$:

$$(6.8) \quad P^{128} = \begin{vmatrix} .390089 & .579037 & .016631 & .014244 \\ .392503 & .586246 & .011831 & .009419 \\ .009465 & .013138 & .487509 & .489888 \\ .011385 & .015999 & .485107 & .487509 \end{vmatrix}$$

$$(6.9) \quad P^{(128)^2} = \begin{vmatrix} .200776 & .298656 & .250286 & .250282 \\ .200782 & .298664 & .250279 & .250275 \\ .200222 & .297829 & .250973 & .250976 \\ .200225 & .297833 & .250970 & .250973 \end{vmatrix}$$

The characteristic roots of P are $\lambda=1, .9996, .95, .92$.

Hence $\lambda^{128} = 1, .95, .0014, 0+ \dots$; $\lambda^{(128)^2} = 1, .0014, 0+, 0+$. That is to say: $P^{128} \approx \pi(1) + \lambda_2^{128} \pi(2)$; while $P^{(128)^2} \approx \pi(1)$.

7. Some Concluding Comments

In the preceding sections, we have analyzed the structure of stable dynamical systems represented by nearly-decomposable matrices. We have seen that such systems may be viewed as composite systems, constructed by the superposition of: (1) a number of highly damped terms, describing the dynamics of separate subsystems within the total system; and (2) the remaining terms, less rapidly damped, describing the dynamics of the interconnections among subsystems.

If such a system starts from a position of disequilibrium, the process of reaching equilibrium may be divided into two phases, corresponding to the decomposition of the dynamical system that has just been indicated. In the first phase, each of the subsystems reaches (or comes very close to) a short-run equilibrium; in the second phase, the entire system moves toward equilibrium, each of the subsystems remaining very close to its short-run equilibrium throughout the process. Hence, in the second phase, the set of variables in each subsystem satisfies (approximately) the Lange-Hicks condition and can therefore be aggregated into a single variable.

Hence, the system variables in the case just described can be represented as a two level hierarchy, with the aggregative variables at the higher level. Now, there is no reason why we need to restrict ourselves to a two-level hierarchy. For in such a hierarchy, each of the subsystem variables at the lower level might be an aggregate of variables at a still lower level of aggregation. The matrix of a

three-level hierarchy, for example, might look something like this:

P_1	Q_1	S_1	S_2
Q_2	P_2	S_3	S_4
P_3	S_2	P_4	Q_3
R_3	R_4	Q_4	P_4

In this matrix, the elements of the submatrices designated as Q 's would be of the first order of smallness, and the elements of the P 's and S 's of the second order of smallness. At the first level of aggregation, there will be four aggregative variables corresponding to the four submatrices along the diagonal, respectively. At the second level of aggregation, there will be two aggregative variables, corresponding to the blocs indicated by broken lines.

To invert a matrix like the one depicted above, we would first invert the matrices P_1 , then the two aggregative matrices derived from

$$\begin{vmatrix} P_1 & Q_1 \\ Q_2 & P_2 \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} P_3 & S_3 \\ Q_4 & P_4 \end{vmatrix}, \text{ respectively,}$$

and finally the second level aggregative matrix.

In ordinary methods of matrix inversion, the number of multiplications increases as the cube of the size of the matrix. In the most favorable case of a two-level hierarchy, using the methods described here, we will need to invert matrices of about size $n^{\frac{1}{2}}$, and there will be about $n^{\frac{1}{2}}$ such matrices. Hence, the number of multiplications will go up approximately as the square of the size of the matrix. But if m , the size of the largest matrix to be inverted at any level of aggregation, remains constant as n increases, then the number of matrices to be inverted

will increase proportionately with n , their size will not increase, and the total number of multiplications will increase only proportionately with n .

It may be objected that decomposable matrices are rare objects, mathematically speaking, and nearly-decomposable matrices almost as much so. For if the elements of a matrix are selected in any ordinary way by a random process, the probability that the matrix will be decomposable is zero. But there is every reason to believe that near-decomposability is a very common characteristic of dynamical systems that exist in the real world. Instantaneous, or nearly instantaneous, action at a distance is rare, and sharp boundaries around subsystems relatively common. We have less assurance that this holds for social phenomena than that it holds for most natural phenomena, but the frequent occurrence of "nearly-diagonal" matrices in input-output analysis argues well for this general approach to aggregation.

The notion that most dynamical systems that are encountered empirically consist of hierarchies of subsystems -- the linkage within the subsystems at each level being stronger than the linkage between subsystems -- is attractive from another standpoint. The probability that a matrix with elements selected independently and at random from a rectangular distribution will have stable roots becomes vanishingly small as the size of the matrix increases (Ashby). If, however, the matrix elements are selected in such a way that most matrices are "nearly diagonal", stability becomes an explicable phenomenon. It is probably no accident that most complex structures we encounter are hierarchical -- or at least have levels.

I should like to call attention to a number of discussions in the economic literature upon which the notion of nearly-decomposable

references is far from complete, since I did not originally approach this problem from the standpoint of aggregation, and became aware of the relation between the two problems only at a relatively late stage in my analysis.

1. I have already pointed out that the argument here may be regarded as a statement of the circumstances under which the Lange-Hicks condition will be satisfied. It can easily be seen from our analysis that if the micro-system is dynamically stable, this will also be true of the aggregated system, since the characteristic roots of the aggregative matrix are also roots of the original matrix. I have found that this stability theorem had been proved earlier in Tamotsu Yokoyama, "A Theory of Composite Commodities," Osaka Economic Review, May, 1952. Yokoyama assumes that the Lange-Hicks condition is satisfied, and derives the stability theorem from this assumption.

2. I have never been satisfied that Goodwin's justification of partial dynamics when two subsystems are unilaterally coupled holds up if the coupling is only "nearly-unilateral." On the basis of the present analysis, I would suggest that the justification for partial equilibrium analysis or partial dynamics should rest on the property of "near-decomposability."

3. The experiments by Horgensorn and Whitin with aggregation in estimating elements of the inverses of Leontief matrices (reported in Input-Output Analysis: an Appraisal) appear to have yielded results consistent with the theory advanced here.

4. Samuelson points out that aggregation of commodities can seek its justification in either of two kinds of principles that, at

first blush, appear rather antithetical to each other. On the one hand (the Lange-Hicks condition), we can aggregate the parts of a subsystem when these are much more closely linked with each other than they are with the rest of the system. On the other hand, we can aggregate a set of variables if each of them is linked with the remainder of the system in just the same way as are the others. Our analysis of near-decomposability shows that the former condition is really a special case of the latter. For if i and j are variables belonging to different subsets of a nearly-decomposable system, then p_{ij} is very small, but $p_{ij}^{(t)}$, for sufficiently large t is almost independent of i . That is to say, the linkage between i and j is negligible in the short run, and satisfies the second condition in the middle run for which it is not negligible.

5. Several persons to whom I have reported orally this method of matrix inversion have conjectured that it may have some relation to the method proposed by Gabriel Kron in a number of his recent publications. Since neither I nor the persons who have mentioned this possibility can understand Kron's papers, I am unable to ascertain whether this is the case. In any event, he does not appear to make explicit use of the characteristic roots or the idempotent components upon which the method developed here depends, nor does he give any indication of the conditions under which aggregation may be expected to give a sufficiently close approximation to the inverse.